M203J Practice Quiz 2 Covering Chapters 13 and 14

Instructions: Answer each question, and when required explain your answer. Your explanation must be clear and complete. You may refer to your book, your notes and your homework papers.

[1] For each of the following kinds of data, choose an appropriate method for graphing the data and an inappropriate method. Indicate which is which, and explain what is inappropriate about the one which is not appropriate. Answers can vary.

(a) Annual profits of a company over a period of years.

Answer: A line graph would be appropriate, in order to emphasize the trend over time. A frequency histogram would not be appropriate; a frequency histogram would make sense in a situation in which we showed the number of years the profits were in various ranges.

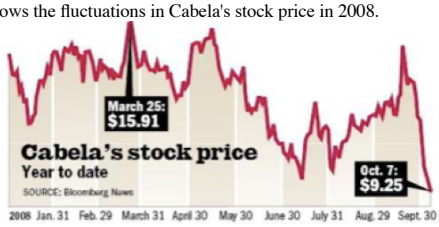
(b) The percentages of families below the poverty level in five different Nebraska cities.

Answer: A bar chart would make sense here, but not a pie chart since the percentages do not describe parts of a whole (they could add up to more than 100%, for example).

(c) The percentages of Nebraska families whose annual incomes are in the following ranges: up to $50,000, above $50,000 but below $100,000, and above $100,000.

Answer: A pie chart would make sense here, since the percentages do describe parts of a whole (they have to add up to 100% in this example). So would a line graph (to emphasize the trend in the percentage as income increases) and a histogram (which is probably what most people would use to graph this data). So all of the types of graphs we discussed are appropriate.

[2] The following line graph shows the fluctuations in Cabela's stock price in 2008.



(a) Discuss what is misleading about this graph. Answers can vary.

Answer: The cropping makes the March 25 data point look much better compared to October 7 than it really is.

(b) Explain how you would redraw this graph so that it is not misleading.

Answer: Start the vertical scale at 0, and include labels on the vertical scale.

[3] Explain how to select a simple random sample of 4 elements from the whole numbers running from 1 to 100, using the random number table handout from class (a copy is included with the quiz). What sample do you get? Explain in enough detail that I can verify that your sample is the one you should have gotten.

Answer: Randomly pick a starting entry in the table, say the entry in row 2 column 4 (which is 64569). Then read down and pick the last two digits of each entry, skipping an entry if it gives a number already chosen. (If the two digits are 00 then that counts as 100.) Here is the simple random sample I get: 69, 7, 38, 8.

[4] A farmer who wants to assess the level of pest infestation in her orchard is considering several different sampling methods. For each of (a) through (d), indicate what sampling technique it corresponds to (choose your answers from among convenience sampling, simple random sampling, systematic sampling or stratified sampling).

(a) Randomly select a tree and then move row by row through the orchard, selecting every 15th tree for inspection: systematic sampling

(b) Pick the first 10 trees near the entrance to the orchard: convenience sampling

(c) Pick a random sample of trees from each of three age ranges: 0 to 3 years old, 3 to 6 years old and 6 or more years old: stratified sampling

(d) Each tree is numbered and a random sample of those numbers is chosen to select the trees to be inspected: simple random sampling

[5] Consider the data 7, 8, 8, 1, 5, 6, 6, 9, 11, 20.

(a) Find the mean of this data.

$(7 + 8 + 8 + 1 + 5 + 6 + 6 + 9 + 11 + 20)/10 = 81/10 = 8.1$

(b) Find the median of this data.

First put the data in order: 1, 5, 6, 6, 7, 8, 8, 9, 11, 20. The median is $(7+8)/2 = 7.5$

(c) (c) Find the first and third quartiles for this data.

1, 5, 6, 6, 7, 8, 8, 9, 11, 20

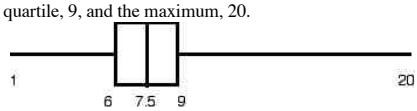
Q_1 = median of lower half of data = 6; Q_3 = median of upper half of data = 9;

(d) Find the range of this data.

the range is $20 – 1 = 19$.

(e) Create and label a box and whisker plot of this data.

The five number summary is the minimum, 1, the first quartile, 6, the median, 7.5, the third quartile, 9, and the maximum, 20.



(f) Find the standard deviation of this data.

the variance is

$[(1-8.1)^2+(5-8.1)^2+(6-8.1)^2+(6-8.1)^2+(7-8.1)^2+(8-8.1)^2+(8-8.1)^2+(9-8.1)^2+(11-8.1)^2+(20-8.1)^2]/10 =$

22.1, so the standard deviation is the square root of this, or about 4.7.

[6] Since 24000 out of 1.1 million fish were tagged, we expect the same fraction fish to be tagged in the second sample of 10300; i.e., we expect $24000*10300/1100000 = 24*103/11 = 2472/11 = 224.7$ which rounds up to 225 fish.